

## Deux moments dans l'histoire du théorème d'algèbre de Ch. F. Sturm\*

**RÉSUMÉ.** — Au départ (1829), le théorème d'algèbre de Ch. F. Sturm fournit un algorithme pour compter le nombre de racines d'un polynôme sur un intervalle réel. Généralisé par A. Tarski (dans les années 1930), il devient un algorithme de décision pour la théorie logique du premier ordre du corps ordonné des nombres réels. L'article suivant considère ces deux moments : la naissance du théorème de Sturm à partir du théorème analogue de J. Fourier ; sa transformation par Tarski qui en révèle la portée logique et inaugure ainsi le renouvellement auquel ce théorème doit encore son actualité.

Diverses questions générales sont abordées par le biais de cette étude : sur les échanges entre Algèbre et Analyse, sur les définitions par conditions nécessaires et suffisantes, sur la nature algébrique des méthodes, sur l'effectivité des procédures, sur la notion d'algorithme.

**SUMMARY.** — *At the outset (1829), Ch. F. Sturm's theorem of algebra provided an algorithm for counting the number of roots of a polynomial which lie within an interval of real numbers. Generalized by A. Tarski during the 1930s, this theorem became a decision algorithm for the first-order theory in logic of the ordered field of the real numbers. In the following article, I consider these two events : the origins of Sturm's theorem, which lay in J. Fourier's similar theorem ; and the transformation of Sturm's theorem by Tarski, which displayed its logical significance and, in this way, ushered in the revival to which this theorem still owes its actuality.*

*From the viewpoint that I take in this study, I take up various general questions : about exchanging algebra and analysis : about definitions in terms of necessary and sufficient conditions ; about the algebraic nature of methods ; about the effectivity of procedures (meaning whatever enables procedures such as finitely recursive procedures to be carried out in a finite number of steps) ; and about the notion of algorithm.*

\* Ce texte est issu de deux communications, la première au Colloque d'Histoire des Mathématiques, CIRM (Marseille-Luminy, juin 1985), sur « Le théorème de Sturm et l'analyse des équations déterminées de J. Fourier » ; la deuxième au Colloque international de Philosophie et d'Histoire des Mathématiques à l'Université autonome de Mexico (décembre 1985), « Histoire et actualité du théorème de Sturm ». Tel qu'il se présente ici, il constitue une étape dans l'élaboration d'un travail plus ample, à paraître en 1989, sous le titre : « Corps et modèles. Aspects de la construction de l'algèbre réelle ».

L'intérêt pour le théorème d'algèbre de Ch. F. Sturm est lié à des questions telles que : Qu'est-ce qu'un *algorithme* ? Qu'est-ce qu'une méthode *effective* ? Qu'est-ce qu'une propriété ou une procédure *algébrique* ? Pourquoi la théorie des équations algébriques s'offre-t-elle comme paradigme pour la *résolubilité* d'un problème ? Or, ces questions, dont l'enjeu est assez général pour tomber dans le champ d'une épistémologie mathématique, sont posées avec assez d'exactitude par la logique mathématique moderne, et en particulier par la *théorie des modèles*. Autant que la théorie de la démonstration, celle-ci a pris en effet le relais de la métamathématique de Hilbert qui, à côté du problème du fondement des mathématiques (que nous mentionnons ici seulement pour mémoire), esquissait ou formulait des interrogations de ce type. A preuve la formulation dans les *Grundzüge der theoretischen Logik* (1928) du fameux problème de la décision qui a marqué toute une époque.

La théorie des modèles, au moins dans l'esprit des grands initiateurs : Alfred Tarski et Abraham Robinson, prétend, par l'étude des langages de théories mathématiques particulières, prises justement pour *modèles*, obtenir des schémas d'énoncés valables pour plusieurs théories à la fois ou, comme on dit, pour toute une classe de modèles. Une procédure, un théorème ne sont donc pas étudiés pour eux-mêmes, pour leur efficacité immédiate. Ils sont repensés dans le cadre d'une métathéorie, qui fait flèche aussi bien de l'information mathématique intrinsèque que de son interprétation logique pour y voir des modèles particuliers de schémas plus généraux. Le gain étant que ces schémas généraux sont passibles d'une description formelle sachant exploiter les ressources d'un langage déterminé *a priori* par un vocabulaire de base et une grammaire.

Le philosophe sera naturellement séduit par l'aspect formel de ce type d'analyse. Le mathématicien pourrait s'intéresser aux retombées techniques qui s'ensuivent le plus souvent. Car il n'est pas indifférent, du point de vue mathématique, de mieux délimiter le domaine de validité d'un énoncé mathématique, de s'apercevoir par exemple que celui-ci est beaucoup plus large que le domaine associé à l'origine et durant le temps d'une première évolution, forcément limitée par le domaine d'origine. Les résultats de la théorie des modèles peuvent ainsi constituer une source de renouvellement et de généralisation de résultats mathématiques.

Cet état de choses, qui milite en faveur d'une coopération entre

mathématiques et logique, et donc aussi en faveur d'une histoire qui tienne compte simultanément d'acquis de ces deux disciplines, est merveilleusement illustré par le cas du théorème d'algèbre de Sturm.

# I — ÉNONCÉ DU THÉORÈME D'ALGÈBRE DE STURM

Il s'agit du théorème présenté par Sturm à l'Académie des Sciences le 25 mai 1829, énoncé et brièvement commenté dans le tome XI (juin 1829) du *Bulletin de Férussac*, dont Sturm était alors rédacteur, exposé complètement et démontré en 1835 dans le « Mémoire sur la résolution des équations numériques » <sup>(1)</sup>.

Ce théorème donne une méthode pour déterminer le nombre de racines réelles, comprises entre deux nombres réels A et B, d'une équation polynomiale  $V = 0$  où V a la forme :

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

Cette méthode consiste à calculer d'abord la dérivée  $V'$ , puis à appliquer l'algorithme d'Euclide pour trouver le p.g.c.d. (plus grand commun diviseur) de V et  $V'$ . On pose  $V' = V_1 Q_1$ , en sorte que  $V = V_1 Q_1 + R_1$ . On dévie alors de l'emploi habituel de l'algorithme d'Euclide en posant  $V_2 = -R_1$ . La raison de ce changement de signe apparaîtra plus bas. On a donc  $V = V_1 Q_1 - V_2$ .

On réitère l'opération avec  $V_1$  et  $V_2$ , pour obtenir :

$$V_1 = V_2 Q_2 - V_3.$$

On recommence avec  $V_2$  et  $V_3$  et on poursuit le processus de division jusqu'à parvenir à un reste  $V_r$  qui, si V et  $V_1$  n'ont pas de racine commune, cas auquel il est toujours possible de se ramener, est une constante numérique (l'indice  $r$  étant majoré par le degré de V). On obtient ainsi le tableau :

$$\begin{aligned} V &= V_1 Q_1 - V_2 \\ V_1 &= V_2 Q_2 - V_3 \\ V_2 &= V_3 Q_3 - V_4 \\ &\dots \dots \dots \\ V_{r-3} &= V_{r-2} Q_{r-2} - V_{r-1} \\ V_{r-2} &= V_{r-1} Q_{r-1} - V_r. \end{aligned}$$

<sup>(1)</sup> In *Mémoires présentés par divers savants à l'Académie [royale] des Sciences de l'Institut de France*, 6 (1835), 271-318. Pour un aperçu de la vie et des travaux de Ch. F. Sturm, voir l'article de P. Speziali dans le *Dictionary of Scientific Biography*, XIII (New York, 1976), 126-132.

La suite des fonctions polynomiales  $V, V_1, V_2, \dots, V_r$  est communément appelée « suite de Sturm », ou « chaîne de Sturm » (l'usage de cette dernière expression remontant à l'exposé du théorème de Sturm fait par Heinrich Weber dans le tome I de son *Lehrbuch der Algebra*, § 91-95) <sup>(2)</sup>.

Etant donné deux nombres réels  $A, B, A < B$ , on écrit la suite  $S_A$  des signes que prennent les fonctions  $V, V_1, V_2, \dots, V_r$ , pour la valeur  $A$  ; et de même la suite analogue  $S_B$ . On compte le nombre  $N_A$  des variations de signe dans  $S_A$ , le nombre  $N_B$  des variations dans  $S_B$ . On démontre que  $N_A - N_B$  est exactement le nombre de racines réelles de  $V = 0$  comprises entre  $A$  et  $B$ .

En son temps, ce théorème fit, à lui tout seul, davantage pour la célébrité de Sturm que tous ses travaux ultérieurs sur les équations différentielles du second ordre, dont l'intérêt ne fut perçu que par la suite (théorie de Sturm-Liouville, à son tour mieux connue aujourd'hui de la grande majorité des mathématiciens). Terquem, par exemple, dans un article des *Nouvelles Annales de Mathématiques* <sup>(3)</sup>, écrit, en 1843, que cette découverte représente un « événement considérable », et qu'elle « produisit une profonde sensation... sur les géomètres de toute classe ». Il ajoute avec une pointe de malice que M. Mayer-Dalembert, « doué d'un grand esprit de spéculation et sachant le prix de toute chose, acquit le droit de faire usage de (cette) invention » et l'inséra en 1832 dans le manuel qu'il publia avec M. Choquet, sous le titre *Traité élémentaire d'algèbre*. Choquet et Mayer-Dalembert furent rapidement imités par Lefébure de Fourcy <sup>(4)</sup> et par Bourdon <sup>(5)</sup> ; puis par A. Serret <sup>(6)</sup> et H. Weber pour ne citer en dernier que les deux auteurs de manuels d'algèbre les plus célèbres dans la deuxième moitié du XIX<sup>e</sup> siècle. En fait, « exigé tacitement pour l'admission à l'école polytechnique », comme le note encore Terquem, ce théorème était vite devenu « obligatoire pour tous les auteurs d'éléments ».

Du côté de la recherche, il ne resta pas non plus lettre morte. Il fut suivi, en effet, par toute une série de mémoires ou d'articles

<sup>(2)</sup> (Braunschweig : Vieweg und Sohn, 1892) ; trad. franç. du t. I (Paris : Gauthier-Villars, 1898).

<sup>(3)</sup> T. II (1843), 97-106 ; t. III (1844), 188-194 ; 209-213 ; 555-565 ; 577-580.

<sup>(4)</sup> *Leçons d'algèbre* (Paris : Bachelier, 1833), chap. XIX, 483-493.

<sup>(5)</sup> *Eléments d'algèbre*, 7<sup>e</sup> éd. (Paris : Bachelier, 1834), chap. VIII, § III, 557-572.

<sup>(6)</sup> *Cours d'algèbre supérieure*, 3<sup>e</sup> éd. (Paris : Bachelier, 1866).

visant à l'améliorer ou à le transformer, de la plume des plus grands mathématiciens comme Sylvester, Cayley, Hermite ou Kronecker, et aussi de moins grands comme Borchardt ou Darboux <sup>(7)</sup>. Il engendra ainsi ce que Sylvester a appelé « un cycle d'idées sturmiennes ».

Ce cycle s'acheva pourtant, sinon avec les travaux de Sylvester <sup>(8)</sup> comme celui-ci voulait bien le croire, du moins avec la difficile théorie des caractéristiques de Kronecker <sup>(9)</sup>.

Après la publication par Liouville des travaux de Galois et l'émergence, en algèbre, d'un mode de pensée structural et de plus en plus orienté vers la constitution de concepts abstraits définis par des systèmes d'axiomes, l'intérêt pour le théorème d'algèbre de Sturm (et pour les questions d'analyse numérique) décru suffisamment pour que celui-ci devînt une matière facultative, reléguée dans l'oubli par les auteurs de manuels.

On peut nuancer ; l'éclipse n'est pas totale. En Allemagne par exemple, la notoriété et la fréquence d'utilisation du *Lehrbuch der Algebra* de H. Weber sont telles que le théorème d'algèbre de Sturm, qui s'y trouve exposé dans un chapitre de quarante pages, ne disparaît pas de l'horizon des chercheurs et des professeurs. En particulier, Artin et Schreier qui en font un usage important en 1926

<sup>(7)</sup> On trouvera une liste de travaux procédant du théorème de Sturm dans G. Mignosi, Teorema di Sturm e sue estensioni, *Rendiconti del Circolo Matematico di Palermo*, 49 (1925), 1-164 ; et une autre liste dans M. Lecat, *L'intermédiaire des mathématiciens*, t. II, 2<sup>e</sup> série (1923), 44-47.

<sup>(8)</sup> Sylvester écrit d'abord deux notes brèves en 1839 et 1841 respectivement. En 1853, il consacre toute la quatrième partie de sa théorie des syzygies à des considérations inspirées par le théorème de Sturm, avec des résultats récemment réactualisés par le développement de la théorie algébrique des formes quadratiques. Voir à ce sujet H. Benis-Sinaceur, communication au XVII<sup>e</sup> Congrès international d'Histoire des Sciences (Berkeley, California, 31 juillet-8 août 1985) ou l'ouvrage annoncé au bas de la p. 99.

Les travaux de Sylvester auxquels nous renvoyons sont reproduits dans *The Collected Mathematical Papers of James Joseph Sylvester* (Cambridge, 1904), t. I, 45-46, 59-60, 396-398, 429-586, 609-619, 620-626, 627-629, 630-633, 634-636, 637-640.

<sup>(9)</sup> Elle est exposée dans plusieurs articles reproduits dans les *Werke* de Kronecker (Leipzig, 1895-1930) : t. I (1895), 175-212, 213-226, 227-234 et 303-348 ; t. II (1897), 37-40, 71-82. Elle est résumée dans l'article rédigé par C. Runge pour l'*Enzyklopädie der Mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, I (1898-1904), 404-448 ; et présentée sommairement par H. Weber, *Traité d'algèbre supérieure* (Paris Gauthier-Villars, 1898), § 101.

Un exposé moderne est donné par H. W. Siegbert, Some historical remarks concerning degree theory, *American Mathematical Monthly*, 88 (1981), 125-139. (Nous remercions M. F. Coste-Roy pour cette référence.)

dans « Algebraische Konstruktion reeller Körper » <sup>(10)</sup> et A. Tarski qui le transforme (vers 1930) en méthode d'élimination des quantificateurs dans son mémoire sur la *Complétude de l'algèbre et de la géométrie élémentaires* <sup>(11)</sup>, se réfèrent à l'exposé de H. Weber. Et le traité de B. L. Van der Waerden <sup>(12)</sup>, qui consacre le dernier chapitre de son premier tome à la théorie d'Artin-Schreier, le consigne naturellement.

Aux Etats-Unis, les mathématiciens sont, à la fin du xix<sup>e</sup> et au début du xx<sup>e</sup> siècle, très liés à leurs maîtres ou confrères européens, allemands en particulier. Certains manuels sur la théorie classique des équations mettent le théorème de Sturm en bonne place. C'est par exemple le cas du *New First Course in the Theory of Equations* de L. Dickson <sup>(13)</sup>. Et l'histoire des mathématiques de Florian Cajori (1<sup>re</sup> éd., 1893) relève à la fois le succès du théorème qui éclipse celui de J. Fourier, et l'étendue de son retentissement : en Angleterre (De Morgan, Sylvester), en Autriche (Andreas von Ettingshausen), en France (A.-L. Cauchy, J. Liouville, J.-M.-C. Duhamel, etc.) <sup>(14)</sup>.

<sup>(10)</sup> *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, 5 (1926), 85-99.

<sup>(11)</sup> Ce mémoire, prêt pour l'édition en 1940 (Paris : Hermann) ne parut qu'en 1967 sous le titre : *The Completeness of Elementary Algebra and Geometry* (Paris : Institut Blaise Pascal) ; trad. franç. dans Tarski, *Logique, sémantique, métamathématique*, t. 2 (Paris : A. Colin, 1974). Entre-temps, une deuxième version en fut publiée aux Etats-Unis sous le titre : *A Decision Method for Elementary Algebra and Geometry* (Berkeley : California, 1948 ; 2<sup>e</sup> éd., 1951).

<sup>(12)</sup> *Moderne Algebra*, 1<sup>re</sup> éd. (Berlin, 1930). On peut remarquer, par contre, que Helmut Hasse, qui publie en 1927-1933 un petit manuel d'algèbre supérieure (*Sammlung Götschen*, Bd. 931-932 ; Berlin : Walter de Gruyter) mentionne dans une note (II, p. 79) la théorie d'Artin-Schreier ; mais son livre, tout entier écrit dans l'esprit de la construction de l'algèbre abstraite, fait un exposé totalement « moderne » de la théorie des équations en commençant par définir les corps, les anneaux, les groupes, les déterminants et en mettant à profit le mémoire de Steinitz paru dans le *Journal de Crelle*, 137 (1909). Dans cette perspective, il n'y a pas de place pour des considérations comme celle de Sturm.

<sup>(13)</sup> (New York, 1932), § 57, 58, 59.

<sup>(14)</sup> A. De Morgan a salué dans le théorème « la solution *théorique* complète d'une difficulté à laquelle s'étaient employées des énergies de tout ordre depuis l'époque de Descartes ». Cité par F. Cajori, *A History of Mathematics* (New York : Macmillan, 1893, 2<sup>e</sup> éd., 1919), 363.

Pour J.-J. Sylvester, se reporter aux références de la n. 8. Donnons les références relatives aux autres auteurs :

— A. von Ettingshausen, Sturm's Regel... nebst einem Beweise derselben, *Z. f. Phys. Math. u. verwandte Wiss.*, t. 7, 444-450.

— A. L. Cauchy, Extrait du mémoire présenté à l'Académie de Turin le 11 octobre 1831 (1832-1833) : *Œuvres*, 2, XV, 262-411. — Extrait d'une lettre sur un mémoire

Mais, en France, la première gloire du théorème de Sturm semble ternir. Sans vouloir établir une stricte chronologie de son maintien ou de son effacement, disons que l'extraordinaire développement en notre siècle de toutes les branches mathématiques fit de cet énoncé singulier un résultat marginal, rarement étudié ou même simplement connu. Prenez, par exemple, dans les années 70-80, un bon étudiant ayant accompli à Paris le cursus normal d'une maîtrise de mathématiques : il a fait de l'algèbre, il connaît un bout de la théorie des équations, il a même fait parfois de l'analyse numérique, sans jamais avoir entendu parler de la méthode de Sturm. Il a sans doute consulté le manuel de R. Godement ou celui de Birkhoff-Mac Lane, il a feuilleté les fascicules d'algèbre de Bourbaki ; nulle part il n'aurait pu trouver une mention, même rapide, de notre théorème. Celui-ci ne semble plus faire partie du savoir vivant. Tout au moins du savoir académique, transmis par l'enseignement.

Les historiens des mathématiques, dont les vues sur le passé dépendent beaucoup des résultats du présent, le mentionnent tout aussi rarement. Le gros livre de Morris Kline, *Mathematical Thought from Ancient to Modern Times* (New York, 1972), qui s'est largement imposé avec une information couvrant deux millénaires et toutes les branches mathématiques, ne souffle mot du procédé de Sturm pour compter et localiser les racines réelles des polynômes. De fait, une périodisation tranchée de la théorie des équations en époque classique ou pré-galoisienne et époque moderne déterminée par Galois, Kronecker, Steinitz, etc., toute légitime qu'elle soit dans son principe, porte « naturellement » à négliger certains résultats (dont le théorème de Sturm). A peu près contemporains, par leur date, de la théorie de Galois, ceux-ci n'ont évidemment pas de raison d'apparaître dans un chapitre sur « L'Algèbre du XVIII<sup>e</sup> siècle » (p. 592-613 de M. Kline). Mais franchement différents par l'esprit, ils ne peuvent non plus figurer dans l'exposé de cette

publié à Turin le 16 juin 1833, *CR*, t. IV, p. 672 (8 mai 1837) ; *Œuvres*, 1, IV, 45-48.

- Note sur un théorème relatif aux racines des équations simultanées, *CR*, t. V, p. 6 (3 juillet 1837) ; *Œuvres*, 1, IV, 81-83. - - Calcul des indices des fonctions, *Journal de l'École polytechnique*, 25<sup>e</sup> cahier (1837) ; *Œuvres*, 2, I, 416-466.

- J. Liouville et Ch. Sturm, Démonstration d'un théorème de Cauchy, *Journal de Liouville*, 1 (1836), 278-289. Et note publiée aux *CR*, t. IV, 1<sup>er</sup> sem. (1837), 720-724.

- J.-M.-C. Duhamel, *Des méthodes dans les sciences de raisonnement* (Paris : Gauthier-Villars, 1867), 2<sup>e</sup> partie, chap. XXVIII, n<sup>o</sup> 231, 257-259.

théorie, ni même dans une liste de résultats précurseurs. Les voilà donc promis à l'indifférence !

Indifférence encore de la part des rédacteurs bourbakistes des *Eléments d'histoire des mathématiques* (Paris, 1960). On ne peut, il est vrai, tout trouver dans des notices si condensées, mais seulement des résultats « très importants », qui *survivent* de façon ou d'autre dans les résultats actuels. Cependant, comment juger de l'importance des résultats actuels ou récents ? Bourbaki, par exemple, s'est peu intéressé à la théorie d'Artin-Schreier. Contrairement à B. L. Van der Waerden, il en donne un exposé très succinct où n'apparaît pas le théorème qui utilise le résultat de Sturm <sup>(15)</sup>. Aussi peut-il écrire, dans une de ses notices historiques, que la question de la séparation des racines d'un polynôme est « un des sujets de prédilection de l'algèbre au XIX<sup>e</sup> siècle » la taxant globalement d'une caducité qui rejaillit en particulier sur le théorème de Sturm <sup>(16)</sup>.

Ainsi, après avoir été porté aux nues par les autorités académiques de son époque, le théorème de Sturm se trouve aujourd'hui, ou jusqu'à une date récente, renvoyé par l'*establishment* mathématique au magasin des antiquités oubliées.

## II — STURM ET L'ANALYSE DES ÉQUATIONS DÉTERMINÉES DE J. FOURIER

Cette situation donne un aperçu des caprices de l'histoire et de l'arbitraire des institutions autant que de la rationalité des évolutions.

D'une part, en effet, il est normal, quoique non obligatoire, qu'un résultat appartenant sans ambiguïté possible à la théorie

<sup>(15)</sup> Il s'agit du théorème d'unicité de la clôture réelle d'un corps ordonné. Voir B. L. Van der Waerden, *Moderne Algebra*, I. La première édition du livre II, *Algèbre* de Bourbaki, chap. VI et VII (Paris : Hermann, 1952), mentionne ce théorème en renvoyant par une note (p. 38) aux pages utiles de la *Moderne Algebra* de B. L. Van der Waerden. La nouvelle édition (*Algèbre*, chap. 4 à 7 (Paris : Masson, 1981)) supprime cette note au profit d'un exercice laissé au soin du lecteur : voir A.VI.24.

<sup>(16)</sup> Celui-ci n'a pas plus de chance dans le récent *Abrégé d'histoire des mathématiques*, publié en 2 volumes sous la direction de J. Dieudonné (Paris : Hermann, 1978 ; nouv. éd. 1986), mais l'*Encyclopaedic Dictionary of Mathematics* publié par Iyanaga (éd. anglaise (Cambridge : MIT, 1977)) mentionne le théorème de Sturm dans son article 296 (XV.5) B, t. II, 932.



classique des équations soit oblitéré par l'offensive de l'algèbre abstraite. D'autre part, le hasard déjoue bien des certitudes, nul n'ayant le moyen de prévoir *a priori* toutes les combinaisons possibles d'idées distinctes, ni, par conséquent, de juger définitivement de l'utilité ou de l'actualité d'un résultat donné.

L'algèbre abstraite se reconnaît pour ancêtre Galois ; pour précurseurs l'école formaliste allemande (Grassmann notamment) et l'école symbolique anglaise (Peacock, De Morgan, Boole), avec une tradition maintenue vivante par Sylvester et Cayley entre autres.

Or Sturm, qui édite certains articles de Galois dans le *Bulletin de Férussac* (entre 1829 et 1831), n'a pas le moindre soupçon du concept de groupe ; ni *a fortiori* du concept de corps. L'idée d'opération abstraite est étrangère à l'ensemble de son œuvre, alors que, à observer la stricte chronologie, elle est déjà présente dans les calculs de Gauss sur les nombres complexes ou sur les classes de formes quadratiques ; et que la théorie va en être faite incessamment dans *l'Ausdehnungslehre* de Grassman (1844) ou les *Lectures on quaternions* de Hamilton (1853 — c'est l'année de la mort de Sturm !). Sturm fait d'ailleurs un usage tout à fait traditionnel des nombres réels sans attention aucune pour les règles formelles de leur calcul. Quant à son théorème d'algèbre, dérivé de ses travaux sur les équations différentielles <sup>(17)</sup>, il ne met en œuvre que des éléments dont la combinaison inédite qu'il représente n'exige aucun élargissement du cadre de la théorie « classique » des équations : ni introduction d'un point de vue nouveau, ni modification, même partielle, du point de vue ancien. Pour ce qui concerne l'algèbre proprement dite, rien dans l'œuvre de Sturm n'annonce ni ne préfigure la mutation que va connaître cette discipline à partir des années 1860 et jusque vers 1930 où le manuel de B. L. Van der Waerden, *Moderne Algebra*, apposera le sceau de la clarté didactique à une somme de recherches dont certaines remontaient au siècle précédent tandis que d'autres avaient à peine trois ou quatre années d'existence !

Sturm fait connaître son théorème en 1829 et le situe explicite-

(17) Cf. la première page de l'extrait de Sturm sur l'intégration d'un système d'équations différentielles linéaires, *Bulletin des Sciences mathématiques, physiques et chimiques* (Férussac), t. XII (1829), art. 196, 313-322 ; et la remarque finale du mémoire Sur les équations différentielles linéaires du second ordre, *Journal de Liouville*, 1 (1836), 186.

ment dans le cadre de « l'analyse algébrique » telle que la comprenait J. Fourier. Il reconnaît franchement sa dette envers son maître et ami. « Je déclare, écrit-il, que j'ai eu pleine connaissance de ceux des travaux inédits <sup>(18)</sup> de M. Fourier qui se rapportent à la résolution des équations, et je saisis cette occasion de lui témoigner la reconnaissance dont ses bontés m'ont pénétré. » <sup>(19)</sup> Puis il ajoute : « C'est en m'appuyant sur les principes qu'il a posés, et en imitant ses démonstrations, que j'ai trouvé les nouveaux théorèmes que je vais énoncer. »

Compte tenu des autres déclarations de Sturm sur son théorème d'algèbre <sup>(20)</sup>, cette phrase résout deux problèmes. Le premier fut agité au XIX<sup>e</sup> siècle et trouve un écho dans l'histoire des mathématiques de Florian Cajori <sup>(21)</sup>. L'impression faite par le théorème de Sturm avait été si forte, en effet, que surgissait parfois la question de sa genèse ; on se demandait comment Sturm en était venu à cette « découverte capitale » <sup>(22)</sup>, à ce « trait de génie » <sup>(23)</sup>. On oscillait entre l'idée qu'il s'agissait d'un « corollaire » des recherches mécaniques de Sturm <sup>(24)</sup>, et celle d'une « illumination subite » <sup>(25)</sup> qu'il aurait eue à la lecture des papiers prêtés par Fourier. Les deux idées sont d'ailleurs émises d'abord par Sturm lui-même comme nous avons voulu le souligner. Et elles ne sont nullement incompatibles, car elles ne sont pas du même ordre. L'une renvoie au *contexte* de la découverte : les travaux sur les équations différentielles, tandis que l'autre en détermine la *texture* : la démonstration.

<sup>(18)</sup> Ces travaux de Fourier, inédits en 1829, seront rassemblés en un volume publié par Navier sous le titre : *Analyse des équations déterminées* (Paris : F. Didot, 1831).

<sup>(19)</sup> *Bulletin des Sciences mathématiques, physiques et chimiques* (Férussac), t. XI (1829), art. 271, 419-422.

<sup>(20)</sup> Voir les textes cités dans les n. 17 et 19.

<sup>(21)</sup> *Op. cit.*, in n. 14, p. 363.

<sup>(22)</sup> Liouville, Eloge funèbre de Sturm cité dans E. Prouhet, Notice sur la vie et les travaux de Ch. Sturm, in Ch. Sturm, *Cours d'analyse de l'Ecole polytechnique* (Paris : Mallet-Bachelier, 1857-1859), t. I.

<sup>(23)</sup> Borchardt, *Nouvelles Annales de Mathématiques*, t. II (1852), 403-406.

<sup>(24)</sup> Liouville, Eloge funèbre de Sturm. Sturm a lui-même (textes cités dans les n. 17 et 19) accrédité ce point de vue qui fut ensuite développé par M. Böcher dans un article souvent cité : The published and unpublished work of Ch. Sturm on algebraic and differential equations, *Bulletin of the American Mathematical Society*, 18 (1911-1912), 1-18.

<sup>(25)</sup> J.-M.-G. Duhamel, *Nouvelles Annales de Mathématiques*, 2<sup>e</sup> série, t. 6 (1867), 427-430 (Réponse à un compte rendu publié par Prouhet dans le même numéro : 234-240). L'idée de Duhamel selon laquelle Sturm a trouvé son théorème par une « espèce de divination », n'est pas si éloignée que l'a pensé F. Cajori de celle de Sylvester qu'il rapporte dans son ouvrage (p. 363).

tration reprend celle du théorème de Fourier, en respecte les principes et en améliore les performances techniques. Aucune des deux ne réduit complètement ce qui, dans une invention, échappe à toute genèse psychologique.

L'autre problème fut posé assez récemment dans le contexte de l'exploitation logico-algébrique du théorème. Prompt, sans doute, à réinterpréter rétrospectivement un résultat en fonction de son utilisation actuelle, on s'est demandé si Sturm n'avait pas déjà eu idée d'une démonstration purement algébrique, libre de toute considération infinitésimale, ou même si, toutes proportions gardées, il n'avait pas un peu anticipé les exigences finitistes du programme de Hilbert <sup>(26)</sup>.

L'étude simultanée du « Mémoire sur la résolution des équations numériques » de Sturm et de l'*Analyse des équations déterminées* (1831) de Fourier montrera qu'il n'en est rien.

## 1 / *Analyse algébrique et analyse infinitésimale selon Fourier*

Le théorème principal de l'*Analyse des équations déterminées* est moins précis que celui de Sturm, puisqu'il donne une majoration du nombre des racines réelles d'un polynôme sur un intervalle, mais ne détermine pas ce nombre de façon exacte <sup>(27)</sup>. Il s'agit en fait, comme son auteur le souligne expressément, d'une généralisation de la règle de Descartes, qui majore le nombre des racines réelles positives d'une équation algébrique par le nombre de variations de signe dans la suite des coefficients de l'équation.

La lecture de l'ouvrage posthume de Fourier n'en est pas moins

<sup>(26)</sup> Voir par exemple G. Kreisel, What have we learnt from Hilbert's second problem ?, *Proc. of Symposia in Pure Math.*, vol. 28 (1976), 102.

<sup>(27)</sup> On identifie communément ce théorème de Fourier et le théorème analogue de Budan. Si cette identité est entière, il faut renoncer à parler du « théorème de Budan-Fourier » car il est facile de montrer l'antériorité des travaux de Fourier. Celui-ci établit son théorème en 1796-1797 dans les leçons à l'Ecole polytechnique ; un résumé de ces leçons a été conservé (voir l'Avertissement de Navier à l'édition posthume de l'*Analyse des équations déterminées*). Mais on a remarqué récemment que cette identification est abusive. A. G. Akritas, par exemple, pense que le théorème de Budan est « différent » quoique équivalent à celui de Fourier. Voir, en particulier, Budan's Theorem and its Consequences, exposé au Séminaire d'Histoire des Mathématiques de l'Université de Nantes, publié dans *Sciences et Techniques en perspective*, n° 4 (1984), 1-13.

instructive. Ne serait-ce que par l'énonciation explicite des *principes* que Sturm reconnaît avoir suivis.

Qu'est-ce que Fourier entend par « analyse algébrique » ?

Cette expression, usuelle depuis la fin du XVIII<sup>e</sup> siècle, reçoit dans l'ouvrage de Fourier deux définitions, l'une par extension, l'autre par compréhension. La première se réduit à énumérer en respectant la chronologie (Préface, p. 2) un certain nombre d'éléments de la tradition mathématique :

- les théorèmes de Viète sur la composition des coefficients,
- la règle de Descartes,
- la règle du parallélogramme de Newton,
- la méthode newtonienne des substitutions successives,
- les recherches de Waring et de Lagrange sur les fonctions invariables des racines et sur l'équation aux différences,
- la théorie des fractions continues telle qu'elle est expliquée dans les ouvrages de Lagrange <sup>(28)</sup>,
- la méthode de D. Bernoulli, déduite des séries récurrentes.

Il s'agit, comme on le voit, d'éléments de la théorie classique des équations, avec en bonne place l'aspect formel et combinatoire des recherches sur les fonctions symétriques des racines.

La seconde définition, qu'on ne trouve pas toute faite dans l'*Analyse des équations déterminées*, est cependant imposée par la philosophie mathématique générale de Fourier et ses idées particulières sur la solution des problèmes laissés en suspens par la résolution algébrique des équations.

En premier lieu, Fourier pense que « les connaissances mathématiques les plus diverses sont toutes de même nature » ; qu'il n'y a donc pas de séparation nette entre les questions dites d'algèbre et les questions dites d'analyse. *Grosso modo*, l'algèbre était alors représentée par la théorie des équations et l'analyse par le calcul infinitésimal. Mais Fourier, pour qui le terme « analyse » a d'ailleurs une résonance qui outrepassa les limites du domaine strictement mathématique <sup>(29)</sup>, appelle « analyse générale » la science des prin-

<sup>(28)</sup> Fourier songe probablement aux Notes et Additions de Lagrange à la traduction française (par Jean III Bernoulli) (Lyon, 1774), des *Eléments d'Algèbre* d'Euler. Ces additions comportent, en effet, un exposé de la théorie des fractions continues.

<sup>(29)</sup> « L'objet de l'Analyse, écrit-il, est un élément préexistant de l'ordre universel. » Ou encore « L'étude approfondie de la nature est... un moyen assuré de former l'Analyse

cipes qui gouvernent toutes les mathématiques et donc aussi bien la théorie des équations (l'algèbre) que le calcul infinitésimal (l'analyse). Dans cette optique unificatrice, il y a une étroite coopération entre l'analyse algébrique et l'analyse infinitésimale.

D'une part en effet, « les questions les plus importantes de la philosophie naturelle, comme celles qui ont pour objet d'exprimer les dernières oscillations des corps, ou la stabilité du système solaire, ou divers mouvements des fluides, ou enfin les lois mathématiques de la chaleur, *exigent une connaissance approfondie des équations* »<sup>(30)</sup>. D'autre part, « on ne peut faire aucun progrès considérable dans la théorie des équations sans quelque usage de l'analyse différentielle, ou, ce qui est la même chose, de la théorie des fluxions » (Préface, p. 3-4). Fourier sait dans quelles difficultés s'empêtrait, au début du XIX<sup>e</sup> siècle, la résolution algébrique des équations et mesure l'intérêt et l'utilité d'une résolution numérique. Dans l'introduction à son ouvrage, il passe en revue et caractérise succinctement tous les travaux connus sur les équations, de Léonard de Pise à Lagrange avec, entre les deux, Cardan, Viète, Wallis, Descartes, Newton, Albert Girard, De Gua, Rolle, Leibniz, Tschirnhaus, Vandermonde et j'en passe ! Si la résolution algébrique fait problème, la résolution numérique « doit être regardée comme une des plus importantes *applications* »<sup>(31)</sup> du calcul différentiel » (Introduction, p. 8, 18 ; exposé synoptique, p. 84-85). Aussi, le lecteur est-il averti de la nécessité de connaître et « les principes » et « l'usage » de ce dernier<sup>(32)</sup>.

Pour des raisons de symétrie autant que par souci d'exactitude, il faut aussi noter le rôle de ce que Fourier appelle « l'analyse des courbes », c'est-à-dire les considérations géométriques sur les tangentes ou sous-tangentes en un point qui correspond à une racine de la fonction polynomiale donnée. L'analyse des courbes concourt, de façon adjuvante ou heuristique, à la solution des problèmes qui ressortissent, tant par leur nature que par la méthode de

elle-même, et d'en découvrir les éléments qu'il nous importe le plus de connaître... » (*Théorie analytique de la chaleur*, Discours préliminaire, in *Œuvres*, I, p. xxii). Et plus loin : « ... l'Analyse suit la même marche dans l'étude de tous les phénomènes ; elle les interprète par le même langage, comme pour attester l'unité et la simplicité du plan de l'univers, et rendre encore plus manifeste cet ordre immuable qui préside à toutes les causes naturelles. »

<sup>(30)</sup> Souligné par nous.

<sup>(31)</sup> Souligné par nous.

<sup>(32)</sup> *Analyse des équations déterminées*, p. 18.

résolution, au calcul algébrique et à l'analyse infinitésimale simultanément. En effet, les relations exprimées par les équations

« n'appartiennent pas seulement à la science abstraite du calcul : elles *existent* <sup>(33)</sup> dans les propriétés des courbes et des surfaces, dans les mouvements des solides et des fluides, dans la distribution de la chaleur, et dans la plupart des phénomènes naturels » <sup>(34)</sup>.

La forme de la courbe, déterminée au moyen du calcul des dérivées successives permet en retour de « connaître distinctement la nature de la fonction primitive », d'où l'on déduit enfin la répartition des racines de l'équation donnée. Ainsi les constructions géométriques tantôt rendent « très sensibles » les résultats du calcul, tantôt aussi permettent de les prévoir <sup>(35)</sup>.

Si la résolution d'une équation dépend de l'application de la théorie des fonctions telle qu'elle s'est constituée à partir du XVIII<sup>e</sup> siècle, ainsi que des techniques du calcul intégral et différentiel, c'est qu'on présuppose toujours que la fonction correspondant à l'équation — qui, du reste, peut être algébrique ou transcendante — est *continue*. Sans l'hypothèse de continuité, au moins locale ou par morceaux, toutes les formules de l'analyse infinitésimale seraient dénuées de sens.

Il y a une correspondance parfaite entre l'application des principes de l'analyse algébrique et l'application des principes de l'analyse infinitésimale. Car, si la première est fondée, comme l'écrit explicitement Fourier <sup>(36)</sup>, « sur la notion générale des variations et des permanences de signe » (instituée par Descartes), on ne peut supposer, par ailleurs, que « les valeurs passent du positif au négatif sans devenir nulles dans l'intervalle » <sup>(37)</sup>. C'est-à-dire qu'appliquer la notion générale des variations et des permanences de signe ne permet de conclure que dans le cas d'une « variation par degrés insensibles » ou continue. « Et lorsque cette condition n'a pas lieu, précise Fourier, rien n'empêche d'examiner séparé-

<sup>(33)</sup> Souligné par nous.

<sup>(34)</sup> *Ibid.*, p. 21. Voir les remarques en ce sens dans *Les voies de l'analyse classique, Critique*, n° 327-328 (août-septembre 1974).

<sup>(35)</sup> Par exemple, sur un intervalle où les racines de  $f(x) = 0$  sont en nombre pair, la construction des tangentes et sous-tangentes à la courbe associée à  $f(x)$  guide la définition d'un critère numérique pour savoir si les racines sont réelles ou imaginaires. Voir *Analyse des équations déterminées*, p. 117 sq.

<sup>(36)</sup> *Ibid.*, exposé synoptique, p. 65.

<sup>(37)</sup> *Ibid.*, p. 59-60.

ment chacune des parties où la continuité subsiste ». L'hypothèse de continuité est donc une condition *sine qua non* de l'analyse des équations numériques. Elle fait basculer la résolution du cadre algébrique, où prime le point de vue *équationnel*, à celui de la *théorie des fonctions*. Il était alors bien naturel que Fourier eût l'idée de substituer, dans l'application de la règle des signes, la suite des fonctions différentielles d'une fonction donnée à la suite des coefficients de l'expression exprimant cette fonction. Cette application nouvelle resitue la règle de Descartes dans un contexte différent de son contexte d'origine.

Sturm, ayant déclaré sa fidélité aux principes de Fourier, n'a pas besoin de revenir explicitement sur l'hypothèse de continuité. Elle est acquise, elle va de soi. Comment supposer alors qu'il ait voulu l'éviter? On peut dire, au contraire, qu'elle a pour lui un degré d'évidence tel qu'il n'a pas, de la première page de son mémoire à la dernière, l'ombre de la moindre réserve.

## 2 / L'impératif d'effectivité

Bien que Fourier connaisse toutes les recherches antérieures sur la théorie des équations, c'est essentiellement par rapport à Lagrange (*Traité sur la résolution des équations numériques de tous les degrés*, 1798) qu'il se définit.

Fourier mesure parfaitement l'enjeu d'une résolution algébrique des équations. Celle-ci consiste

« à assigner pour une équation proposée d'un degré quelconque un nombre limité d'opérations, *tellement ordonnées que le résultat de la dernière fût une des racines*, en n'admettant au nombre de ces opérations à effectuer que les règles élémentaires du calcul (+, ×, :, —) et celles qui sont indiquées par les signes radicaux » (38).

Comme tout mathématicien intéressé par la théorie des équations, Fourier sait le prix d'un *algorithme* de résolution générale, qu'il décrit comme une « opération régulière et facile dont la marche est toujours la même » (39). Malheureusement, « l'analogie du second degré est trop incomplète » en ce qui concerne les degrés supérieurs.

(38) *Ibid.*, p. 13, mais c'est nous qui soulignons la périphrase par laquelle Fourier vise la notion d'algorithme.

(39) *Ibid.*, p. 16.

Fourier sait l'impossibilité de la résolution algébrique générale. Sans s'arrêter à une démonstration « régulière », il en souligne l'inutilité *prévisible* <sup>(40)</sup> et insiste sur l'utilité d'une résolution *numérique*.

Celle-ci a d'ailleurs l'avantage de fournir non seulement un procédé de résolution mais aussi des solutions effectives, du moins si le procédé est d'application *facile*. Celui de Lagrange, par exemple, l'équation aux carrés des différences, irréprochable théoriquement, est impraticable <sup>(41)</sup> ! Alors que les difficultés attachées à une résolution algébrique générale conduisaient à dissocier deux problèmes : celui, prioritaire, de *l'existence* d'une solution par radicaux, et celui du *calcul* de cette solution quand elle existe <sup>(42)</sup>, Fourier maintient indissociées existence de principe et existence effective. Par exemple, même les formules connues pour la résolution algébrique des équations du troisième et du quatrième degré le laissent insatisfait. Car elles comportent des calculs trop complexes pour conduire rapidement à « la connaissance effective des racines » ; elles s'écartent du « but réel » qui est de « connaître en nombres, ou en une suite de monômes » ces racines.

Ce parti pris évident de *calculabilité*, Fourier l'a constamment

<sup>(40)</sup> *Ibid.*, p. 15. Rappelons que le fameux mémoire d'Abel sur l'impossibilité de résoudre algébriquement les équations d'un degré supérieur au quatrième fut republié dans le premier volume du *Journal de Crelle*, en 1826. La même année, un compte rendu en est donné dans le *Bulletin de Férussac* (t. 6, n° 207, 347-354), avec une note sur les résultats antérieurs de Ruffini et le rapport favorable qu'en fit Cauchy devant l'Académie des Sciences. *A posteriori*, on peut naturellement méditer sur la signification de l'adjectif « régulière » sous la plume de Fourier !

<sup>(41)</sup> En ce qui concerne la résolution numérique, Fourier fait le point de la façon suivante (*ibid.*, art. 35 du liv. II) : « La difficulté de distinguer le cas des deux racines imaginaires du cas des deux racines réelles est le point le plus important de l'analyse des équations ; elle exige une méthode propre fondée sur le calcul des limites entre lesquelles les racines sont comprises. Les recherches de Rolle, celles de De Gua, n'ont pu conduire à la résolution numérique des équations, parce qu'elles manquaient d'un caractère spécial pour distinguer les racines imaginaires. Le calcul de l'équation aux carrés des différences a résolu pour la première fois cette singulière difficulté ; mais comme on l'a remarqué depuis longtemps, la solution est purement théorique, et les tentatives que l'on a faites pour la perfectionner... ont été presque entièrement infructueuses. Il était donc nécessaire de traiter la question d'une manière différente... non moins exacte et d'une application incomparablement plus facile. » Sturm a repris le jugement de Fourier sur Lagrange ; voir le premier paragraphe de son mémoire de 1835.

<sup>(42)</sup> Lagrange, par exemple, en distinguant nettement la résolution algébrique des autres types de résolution, conçoit que la première n'obéisse pas forcément à l'impératif d'effectivité. Voir son introduction du *Traité de la résolution des équations numériques de tous les degrés*, p. 16 de l'édition Serret (1879).



défendu dans sa pratique de l'analyse différentielle (résolution de l'équation différentielle correspondant à la propagation de la chaleur). L'évolution, au XIX<sup>e</sup> siècle, de la théorie des équations différentielles qui devient (avec Sturm précisément, puis Liouville) une théorie abstraite, davantage orientée vers la détermination théorique de l'existence et de la forme des solutions intégrales que vers le calcul d'intégrales particulières, l'évolution de la théorie des fonctions qui formalise ses concepts fondateurs, l'évolution de l'algèbre enfin qui cristallise ses procédés en concepts plutôt que de les développer en algorithmes, tout cela concourt à voiler provisoirement l'intérêt *théorique* des méthodes effectives. Kronecker, notamment influencé par le théorème de Sturm qu'il connaît bien, allait certes se faire le grand champion des méthodes effectives. Mais ses justifications auront une allure métaphysique. C'est Hilbert, qui en posant le problème de la décidabilité, révélera la dimension métamathématique de ces méthodes. Le développement de la théorie des fonctions récursives, d'une part, la mise au point d'algorithmes de décision pour des théories mathématiques ou logiques particulières, de l'autre, restitueront sur un autre plan le parti pris de calculabilité. En transformant le théorème de Sturm en algorithme de décision pour l'algèbre et la géométrie élémentaires, Tarski se basera sur son caractère effectif, propriété exigée explicitement par les principes de Fourier que Sturm a faits siens.

### III — LA DÉMONSTRATION DE STURM

#### 1 / *Les principes de Fourier*

En adoptant les principes de Fourier, Sturm combine naturellement calcul algébrique et analyse infinitésimale dans la démonstration de son théorème. Son mémoire de 1835 dresse un argument où l'on peut distinguer des hypothèses générales, qui en délimitent le cadre, et des méthodes plus particulières, qui en déterminent la spécificité. Font partie des présupposés généraux l'hypothèse de continuité et la méthode d'inspection des variations et des permanences de signe. L'une ressortit, on l'a vu, à l'analyse infinitésimale, l'autre au calcul algébrique ; les deux étaient déjà employées par Fourier. Les méthodes spécifiques sont l'algorithme

d'Euclide appliqué, de façon modifiée <sup>(43)</sup>, aux polynômes ; le théorème des valeurs intermédiaires selon lequel toute fonction continue qui change de signe sur un intervalle réel s'annule nécessairement entre les bornes de cet intervalle (avec une réciproque du style : toute fonction continue qui s'annule pour une valeur réelle  $\alpha$  qui n'annule pas sa dérivée change de signe sur un intervalle contenant  $\alpha$ ), et le théorème de Taylor. Ces trois derniers théorèmes, qui présupposent la notion de continuité, sont évidemment, dans la forme où ils apparaissent tant chez Fourier que chez Sturm, du ressort de l'analyse infinitésimale. L'algorithme d'Euclide, qui, appliqué aux polynômes, est souvent appelé par les auteurs allemands algorithme d'Euler, y échappe. Or, il constitue dans le théorème de Sturm, le seul élément absolument nouveau par rapport au théorème de Fourier.

Cette dissection de la démonstration du théorème d'algèbre de Sturm illustre assez la fidélité aux principes de Fourier qui prône, dans un esprit anti-lagrangien, le recours aux méthodes de l'analyse différentielle.

Une analyse détaillée du mémoire de Sturm, qui ne saurait être menée dans les limites de cet exposé, montrerait la complémentarité exacte des arguments algébriques et des arguments infinitésimaux, avec, pour ces derniers, une prééminence due à ce que tous les raisonnements présupposent l'hypothèse de continuité. En fait, il ne suffit pas de démêler, dans la démonstration de Sturm, les arguments algébriques d'avec les arguments infinitésimaux, il faut encore voir comment ils se conjuguent et par quoi passe, en définitive, leur efficacité.

Prenons la méthode d'inspection des variations ou des permanences de signe. Elle est, depuis Descartes au moins, d'un large usage dans toutes les questions relatives aux racines réelles des équations. Comme elle passe pour être un outil algébrique, elle paraît irréprochable du point de vue de ceux qui répugnent aux notions infinitésimales. L'étude simultanée de son emploi par Lagrange, Fourier et Sturm montre, cependant, que son statut n'est, en fait, pas si univoque qu'il le paraît. Statut et mode d'emploi sont liés. Voyons de plus près.

<sup>(43)</sup> Comme le note Lindsay Childs, *A Concrete Introduction to Higher Algebra* (New York-Berlin : Springer-Verlag, 1979) chap. 7, il s'agit d'une « application intelligente de l'algorithme d'Euclide ».

Une des formes d'utilisation principales de cette méthode apparaît dans la corrélation entre racine réelle et changement de signe d'une équation. Or, cette corrélation peut être formulée dans des registres mathématiques divers.

Lagrange s'en tient, pour sa part, à une version purement algébrique, représentée par les théorèmes I et II du chapitre I de son *Traité de la résolution des équations numériques* (1798) <sup>(44)</sup>. Dans la démonstration de ces deux théorèmes « qui sont la base de toute la théorie des équations », Lagrange veut éviter le recours à « la théorie des lignes courbes » et s'en tenir à « la nature même » des *équations*. Cependant il s'y reprend à deux fois (chap. I et n. 1) sans parvenir à un résultat satisfaisant.

Après cet échec, Fourier, on l'a dit, opte franchement pour le recours à « la théorie des lignes courbes », et justifie ce recours par les principes que nous avons évoqués.

Sturm, à son tour, fonde son analyse des équations différentielles autant que son théorème d'algèbre sur ladite corrélation en lui donnant la forme suivante.

D'une part, il admet comme vérité évidente que toute fonction continue, qui change de signe entre deux nombres réels A et B, s'annule au moins une fois dans l'intervalle (A, B) <sup>(45)</sup>. D'autre part et réciproquement, il énonce et démontre non en toute généralité mais dans le contexte de son argumentation, que si une fonction continue s'annule pour une valeur  $\alpha$  qui n'annule pas sa

<sup>44)</sup> Rappelons le texte de Lagrange : « Théorème I : Si l'on a une équation quelconque, et que l'on connaisse deux nombres tels qu'étant substitués successivement à la place de l'inconnue de cette équation, ils donnent des résultats de signes contraires, l'équation aura nécessairement au moins une racine réelle dont la valeur sera entre ces deux nombres.

« Théorème II : Si dans une équation quelconque, qui a une ou plusieurs racines réelles et inégales, on substitue successivement à la place de l'inconnue deux nombres dont l'un soit plus grand et l'autre soit plus petit que l'une de ces racines et qui diffèrent en même temps l'un de l'autre d'une quantité moindre que la différence entre cette racine et chacune des autres racines réelles de l'équation, ces deux substitutions donneront nécessairement deux résultats de signes contraires. »

<sup>45)</sup> Mémoire sur les équations différentielles linéaires du second ordre, § XIV, *Journal de Liouville*, I (1836), 132 ; Mémoire sur une classe d'équations à différences partielles, § XXV, *ibid.*, 432 ; Mémoire sur la résolution des équations numériques, § 3 et 4, *Mémoires présentés par divers savans à l'Acad. des Sciences de l'Inst. de France*, VI (1835).

Ce résultat fut démontré par Bolzano en 1817 (*Rein analytischer Beweis...*) et par Cauchy en 1821 (*Cours d'analyse algébrique*).

dérivée, alors elle change de signe sur un intervalle contenant  $\alpha$  <sup>(46)</sup>.

Il ne s'agit plus d'*équation*, mais de *fonction*, et de fonction continue ainsi que de dérivée. De Lagrange à Sturm la formulation a totalement changé ; les concepts et donc les méthodes sont tout autres. Comme chez Fourier, l'inspection des variations ou permanences de signe se trouve donc immergée dans un contexte qui emprunte beaucoup à la théorie des fonctions et au calcul infinitésimal. Son efficacité présuppose de façon essentielle le concept de continuité.

Sturm ne présente ni ne discute explicitement les réquisits de son argumentation comme nous venons de les schématiser. Cependant il est impossible de ne pas les y voir, pourvu qu'on lise le texte de son mémoire. Du reste, ne s'agit-il pas là de présupposés très généraux dans le contexte de l'analyse algébrique du début du XIX<sup>e</sup> siècle <sup>(47)</sup> ? Et ces présupposés n'ont-ils pas été énoncés par Fourier sous la forme explicite de principes, auxquels Sturm a déclaré son adhésion ?

## 2 / *L'esprit d'abstraction de Sturm*

On peut faire, cependant, une lecture plus distanciée du texte de Sturm. Après tout, le succès d'un résultat se mesure souvent à l'écart entre les idées qu'il inspire et celles dont il est issu !

Or, il y a une sorte d'effet « objectif » du texte de Sturm, produit à moitié par des qualités intrinsèques et à moitié par une projection rétrospective, qui en majore l'aspect algébrique au détriment de l'aspect analytique.

L'idée d'éviter les arguments géométrico-analytiques de continuité n'est ni explicitement défendue ni implicitement soutenue par Sturm. D'ailleurs, c'est seulement à la fin du XIX<sup>e</sup> siècle qu'elle devient une des idées mathématiques dominantes (Dedekind, Hilbert en particulier). Ce qui est vrai c'est qu'il y a, en dépit de l'identité des principes, une différence générale de style entre Fourier

<sup>(46)</sup> Sur les équations différentielles linéaires, § II ; Mémoire sur la résolution des équations numériques, § 4.

<sup>(47)</sup> Voir le *Cours d'Analyse algébrique* de Cauchy et le *Cours d'Algèbre supérieure* de A. Serret.

et Sturm, et une différence plus particulière au théorème d'algèbre, qui facilite sa récupération *a posteriori* par les amateurs de méthodes algébriques.

*a /* Nous n'insisterons ici qu'en passant sur le style général du travail de Sturm grâce auquel on doit la première théorie abstraite des équations différentielles. Quelle que soit la nature de l'équation, différentielle ou algébrique, Sturm attaque le problème de sa résolution en cherchant à déterminer le nombre de changements de signes sur un intervalle réel de certaines fonctions. Cette méthode, il ne l'emploie pas seulement dans le cas où le calcul direct des racines (réelles), est difficile ou impossible. Pour lui, elle constitue de toute façon un préalable. Car elle permet de déceler *a priori* les *propriétés*, sinon les *valeurs*, des fonctions intégrales qu'on ne sait pas calculer. Comme dit Sturm, elle donne sinon la solution effective, du moins « la forme » de la solution, c'est-à-dire l'ensemble des propriétés prévisibles à partir d'une analyse de la *forme* de l'équation proposée.

Sturm se détourne ainsi des développements en sinus et cosinus d'arcs multiples de Fourier et tâche de « reconnaître les propriétés caractéristiques » des fonctions et les sinuosités des courbes intégrales sur « la seule considération des équations différentielles » (48). Ce faisant, il applique à la résolution des équations différentielles une idée caractéristique de l'algèbre traditionnelle des polynômes : le nombre et la nature des racines fournit une caractérisation antérieure à tout calcul effectif. C'est ce transfert de méthode qui produit la mutation bien connue dans l'histoire des équations différentielles et représentée par ce qui est appelé la « théorie de Sturm-Liouville ».

Cette mutation relève, chez Sturm, de la même disposition que celle par laquelle on va le voir définir les « suites de Sturm ». Elle exprime également la conviction — alors peu banale — qu'il est possible d'obtenir des énoncés mathématiques intéressants sur des objets numériquement indéterminés. Et qu'il est souvent plus facile et plus rentable de saisir globalement des ensembles plutôt que des éléments particuliers : la forme de l'ensemble des solutions d'une équation différentielle donne, à moindres frais, plus d'infor-

<sup>48)</sup> Voir, par exemple, l'introduction du « Mémoire sur les équations différentielles linéaires du second ordre ».

mation que le calcul d'une ou plusieurs solutions particulières ; de même, la forme d'une « suite de Sturm » va résoudre un problème sur lequel les mathématiciens butaient depuis longtemps.

*b* / C'était en effet, comme le souligne Fourier à tant de reprises <sup>(49)</sup>, une question bien difficile que celle de la nature, réelle ou imaginaire, des racines d'une équation algébrique.

Rappelons, par exemple, que la règle de Descartes ne nous dit pas s'il y a des racines réelles, mais seulement, au cas où il y en a, que leur nombre ne peut dépasser celui des variations de signe de la suite des coefficients de l'équation proposée. Lagrange, qui avait théoriquement résolu le problème par la formation de l'équation dont les racines sont les carrés des différences entre les racines de la proposée, doit cependant se restreindre au cinquième degré ou faire une hypothèse particulière sur le nombre des racines imaginaires <sup>(50)</sup>. Fourier a l'idée de combiner la règle de Descartes avec le théorème de Rolle en dénombrant les variations de signe de la suite des dérivées successives de la proposée. Mais, comme Descartes, il n'obtient qu'un majorant du nombre de racines réelles, à moins que l'équation considérée n'ait toutes ses racines réelles, et le risque n'est pas supprimé de chercher des racines dans des intervalles où il n'y en a pas.

Connaître le nombre exact de racines réelles d'une équation algébrique, c'est connaître du même coup le nombre de ses racines imaginaires, et le problème de la nature des racines se trouve alors résolu. Sturm est frappé par « l'incertitude » dans laquelle nous laisse même le théorème de Fourier, comme il ressort des premières lignes de son mémoire sur la résolution des équations numériques. On peut penser que, ayant eu l'occasion « d'étudier à loisir » les papiers personnels de Fourier sur la question, il a dû remarquer aussi la longueur et la multiplicité des procédés de celui-ci <sup>(51)</sup>. Dans le cas général en effet, le théorème de Fourier doit obligatoirement être assorti de règles complémentaires si nombreuses qu'on se trouve pris dans les entrelacs d'une sinieuse analyse de cas. La

<sup>(49)</sup> *Analyse des équations déterminées*, Introd. p. 10, 3<sup>e</sup> al. ; Exposé synoptique, p. 29, 2<sup>e</sup> al. ; liv. II, art. 35.

<sup>(50)</sup> *Traité de la résolution des équations numériques de tous les degrés*, chap. V, art. II, n<sup>o</sup> 36.

<sup>(51)</sup> *Analyse des équations déterminées*, p. 31 à 33 ; ou liv. I, § 20 à 38. Le § 38, qui est un résumé des différents cas ou étapes, couvre à lui seul trois pages et demie.

méthode d'inspection des variations de signe ne peut être appliquée de façon *uniforme*, ni surtout *univoque*.

La principale entrave à cela est, comme l'a remarqué Fourier lui-même, que la diminution du nombre de variations de signe ne correspond pas forcément à une racine réelle de l'équation  $f(x) = 0$ . La situation est ambiguë ; sauf dans un seul cas, celui d'une fonction telle que si  $f(a) = f^{(i)}(a) = 0$  pour un certain nombre réel  $a$ , alors  $f^{(i-1)}(a)$  et  $f^{(i+1)}(a)$  sont de signes opposés. La particularité de ce cas s'oppose à un raisonnement général.

Que fait Sturm devant cet obstacle majeur ? Par un artifice dont on peut trouver d'autres exemples dans l'histoire des mathématiques il transforme l'obstacle en levier. Du cas isolé il fait un cas typique. C'est-à-dire qu'il prend la condition restrictive justement comme critère de définition, et pose *a priori* que les fonctions, auxquelles s'appliquera (de façon vraiment efficace parce que *univoque*) la méthode d'inspection des signes, devront satisfaire la condition isolée par Fourier dans un cas particulier. A celle-ci seront associées trois autres conditions mais d'importance moindre <sup>(52)</sup>. Si bien qu'on arrive à une définition de type axiomatique (avant la lettre) qui ne désigne pas un objet particulier, telle suite particulière de fonctions, mais détermine un ensemble de suites (les « suites de Sturm ») par la donnée *a priori* d'un petit nombre de propriétés caractéristiques. Sturm est parfaitement conscient du caractère de sa démarche. Il expose d'abord complètement sa démonstration sur le mode naïf : il construit explicitement une « suite de Sturm » particulière et montre qu'elle fait l'affaire. Puis, sur un mode plus savant (plus abstrait), il dégage et énumère les propriétés nécessaires et suffisantes qui définissent des suites du même type, et souligne ainsi que la conjonction de ces propriétés détermine non pas une suite unique mais une infinité.

Sturm nous donne ainsi un très bel exemple du processus par lequel une définition abstraite est forgée dans l'étude d'un cas particulier. En tranchant dans le vif, en se débarrassant d'entrée de jeu de tous les cas ambigus pour ne s'intéresser qu'au seul cas univoque, en se donnant artificiellement les conditions qui gouvernent ce cas, son théorème atteint une clarté, une simplicité

<sup>52)</sup> Nos 9, 10 et 11 du mémoire de Sturm. Pour une formulation générale de ces conditions, voir l'exposé de H. Weber.

et une généralité qui manquaient à celui de Fourier. Grâce à lui, la méthode d'inspection des variations ou permanences de signe, à laquelle Fourier avait donné une portée bien plus générale qu'elle n'avait chez Descartes, Rolle ou Lagrange, devient enfin un outil sûr : elle fonctionne dans des conditions telles que toute diminution des variations de signe correspond *exactement* à l'existence d'une racine réelle de l'équation proposée. On n'a plus besoin alors de la « considération des lignes courbes », ni, en particulier, de l'étude, indispensable pour Fourier <sup>(53)</sup>, du rapport numérique associé à la sous-tangente en un point à la courbe d'équation  $f(x)$  pour savoir si on a affaire à deux racines réelles ou à deux racines imaginaires.

### 3 / L'utilisation de l'algorithme d'Euclide

Ainsi les principes mis à part, Sturm ne reprend pas *en pratique* tous les raisonnements courants de l'analyse infinitésimale. Cela est d'autant moins nécessaire qu'il laisse tomber la suite des dérivées sur laquelle Fourier avait bâti son théorème, et lui substitue une suite construite par un procédé algébrique classique : l'algorithme d'Euclide. Nous avons déjà dit que c'est là l'innovation majeure par rapport au théorème de Fourier. Innovation aussi par l'adaptation du procédé classique au problème résolu ; à chaque division le signe du reste est changé et ce changement, « inutile si l'on n'avait pour but que de trouver le plus grand commun diviseur, est nécessaire à (cette) théorie ». Nécessaire en effet pour produire la fameuse condition de définition d'une « suite de Sturm ». On voit bien, encore une fois, que Sturm procède artificiellement, c'est-à-dire invente.

L'algorithme d'Euclide modifié ne sert pas, dans le mémoire de 1835, à définir une « suite de Sturm » en général. Pour cela, nous l'avons dit, il n'est besoin que d'énoncer les conditions nécessaires et suffisantes d'existence de telles suites. Il sert à construire *explicitement* une suite particulière, grâce à laquelle, d'ailleurs, Sturm démontre son théorème. Mais il joue un rôle certainement décisif pour la postérité du théorème. Peut-être a-t-il suggéré à Sylvester ou Hermite, par exemple, la possibilité d'un recadrage théorique

<sup>(53)</sup> En particulier liv. I, § 24.



du théorème. Si, au lieu de tout baser sur les concepts de fonction et de continuité, on part du concept algébrique de polynôme, suite finie de nombres appelés coefficients, on peut, avec la construction algorithmique de la suite de Sturm, raisonner d'un bout à l'autre dans un cadre purement algébrique, comme c'était l'ambition de Lagrange que Fourier croyait irréalisable. Si, au lieu de considérer une seule équation à une seule inconnue, on envisage plus généralement des systèmes de  $n$  équations à  $n$  inconnues, alors un lien s'établit entre le théorème de Sturm et des résultats de la théorie de l'élimination algébrique, qui met sur la voie d'une déduction algébrique d'un résultat apparenté à celui de Sturm <sup>(54)</sup>.

#### IV — LE THÉORÈME DE STURM GÉNÉRALISÉ PAR TARSKI

Ce n'est pourtant pas le résultat de Sylvester mais bien celui de Sturm qui attire, au début du  $xx^e$  siècle, l'attention des mathématiciens engagés dans le développement de l'algèbre abstraite comme Artin et Schreier, ou des logiciens interpellés par le problème de la décision posé par Hilbert comme Tarski. L'actualité du théorème de Sturm est surtout liée à son utilisation par Tarski pour démontrer la complétude de l'algèbre et de la géométrie élémentaires <sup>(55)</sup>. Tarski en conçoit une version encore plus générale que celle rendue familière par les travaux d'Hermite et de Sylvester <sup>(56)</sup>. Il envisage, en effet, des systèmes pouvant comprendre  $p$  équations à  $n$  inconnues (pour  $n \neq p$ ) et  $q$  inéquations à  $n$  inconnues. Comme il l'écrit dans la note 12 de la deuxième

<sup>(54)</sup> Nous laissons ici de côté cet aspect, abordé dans notre communication au XVII<sup>e</sup> Congrès international d'Histoire des Sciences (Berkeley, Californie, juillet-août 1985).

<sup>(55)</sup> Ce résultat se trouve dans un mémoire dont on connaît deux versions, formellement assez différentes (voir n. 11 *supra*). Nous nous référons beaucoup à la première, dont le texte commente de façon répétée la nature des services rendus par le théorème de Sturm. Mais la seconde version est assortie d'une très longue note (n. 12) où Tarski reformule le théorème en soulignant son lien à ses propres théorèmes d'élimination des quantificateurs.

<sup>(56)</sup> Pour Sylvester se reporter à la n. 8. Pour Hermite, voir Sur l'extension du théorème de M. Sturm à un système d'équations simultanées, dans les *Œuvres* (Paris : Gauthier-Villars, 1905-1917), t. 1, 281-283 et III, 1-34. Et les Remarques sur le théorème de M. Sturm, *ibid.*, t. 1, 184-287.

version de son mémoire, le théorème de Sturm fournit « un critère de résolubilité dans  $\mathbf{R}$ , qui implique aussi bien des inéquations que des équations ».

### 1 / Un critère « purement algébrique »

Ce qui frappe dans le mémoire de Tarski eu égard au théorème de Sturm, c'est que celui-ci est présenté comme un théorème algébrique. Tarski fait remarquer que, pour les besoins de sa démonstration, « il est d'une importance fondamentale qu'il soit possible de construire pour la fonction donnée  $f(x)$ ... les fonctions (de la chaîne de Sturm) à l'aide de moyens purement algébriques » <sup>(57)</sup>. C'est dire que Tarski est surtout intéressé, dans le théorème de Sturm, par l'emploi de l'algorithme d'Euclide <sup>(58)</sup> qui permet de construire « une suite finie de fonctions rationnelles (de polynômes au sens algébrique) ». Les composantes non algébriques de la démonstration originale : le recours aux idées de variation infiniment petite ou de continuité, Tarski n'y prête aucune attention et les efface totalement, pour ainsi dire, de *son interprétation* du théorème de Sturm.

C'est que, d'une part, Tarski raisonne sur des polynômes et non sur des fonctions continues. Aussi n'a-t-il pas besoin de la version analytique du théorème des valeurs intermédiaires, mais seulement de ce qu'il en considère comme un « cas particulier », celui où on en restreint l'application à des polynômes. La formulation de cet énoncé plus particulier se fait en termes algébriques qui décrivent des *conditions de signe*. C'est l'axiome XVII' de son système d'algèbre des nombres réels <sup>(59)</sup>.

D'autre part, étant donné l'usage qu'il veut faire du théorème

<sup>(57)</sup> La complétude de l'algèbre et la géométrie élémentaires, dans *Logique, sémantique, métamathématique, 1923-1944* (Paris : Armand Colin, 1974), t. 2, p. 219.

<sup>(58)</sup> Pour Tarski, algorithme d'Euclide et théorème de Sturm sont liés de façon essentielle, en dépit de la possibilité de construire des chaînes de Sturm par d'autres moyens. Dans l'introduction à la deuxième version de son mémoire, s'agissant de donner des exemples mathématiques connus de ce qu'on appelle une « méthode de décision », il cite exclusivement l'algorithme d'Euclide (appliqué aux entiers naturels positifs et aux polynômes à coefficients constants) et le théorème de Sturm. Or, on peut penser encore à l'algorithme qui sert à calculer la racine carrée d'un nombre naturel (mais il s'agit d'un algorithme infini, *i.e.* qui peut être indéfiniment continué).

<sup>(59)</sup> « La complétude... », I.12, p. 213-215.

de Sturm : prouver son *lemme d'élimination des quantificateurs* <sup>(60)</sup> pour la théorie élémentaire du corps ordonné des nombres réels <sup>(61)</sup>, il importe à Tarski de fournir un critère purement algébrique <sup>(62)</sup> de l'existence d'une solution réelle pour un système donné d'équations et d'inéquations polynomiales. L'existence d'un tel critère algébrique est en effet l'indice mathématique de l'*existence* d'une procédure (métamathématique) d'élimination des quantificateurs. Le théorème de Sturm fournit bien, comme Tarski le souligne de façon répétée, ce type de critère. Cependant il s'agit évidemment pour Tarski, non du théorème dans sa facture originale, mais de « la méthode de Sturm généralisée » <sup>(63)</sup> qu'il a formulée et redémontrée dans le cadre logico-algébrique qu'il a lui-même tracé. Tarski n'avait pas comme premier souci celui de l'exactitude historique ! Mais Tarski infidèle, c'est Tarski créateur. Il fait un libre usage d'un théorème connu <sup>(64)</sup> et le transforme pour ainsi dire spontanément, en sorte à l'adapter à ses propres buts <sup>(65)</sup>.

Sans entrer dans des détails dont la technicité nous entraînerait

<sup>(60)</sup> Il convient de distinguer le lemme d'élimination des quantificateurs de l'algorithme correspondant. Si le premier est une trouvaille lumineuse, le second est loin d'être aussi simple en pratique (sur une machine) qu'il l'est en théorie. Des travaux récents sont ainsi consacrés à simplifier l'algorithme de Tarski, c'est-à-dire à réduire le temps utile à un ordinateur pour son exécution.

<sup>(61)</sup> « La complétude... », lemme 2.3, p. 217. Y correspondent les théorèmes 27, 29, 31 de *A Decision Method for Elementary Algebra and Geometry*, 2<sup>e</sup> éd. (Berkeley : University of California Press, 1951).

Tarski démontre que tout énoncé existentiel de la forme  $\exists x A(x)$ , où  $x$  représente un nombre réel quelconque et  $A$  une conjonction d'équations et d'inéquations algébriques, est équivalent à un énoncé ne comportant plus aucun quantificateur. Élémentaire se dit d'un énoncé ou d'un ensemble d'énoncés qui peuvent être formulés dans le langage logique du calcul des prédicats du premier ordre (les seules variables quantifiées étant des variables d'individus).

<sup>(62)</sup> « La complétude... », p. 218 : « Le contenu mathématique du lemme 2.3 (d'élimination des quantificateurs) se réduit à ceci. Il est possible, en accord avec ce lemme, de fournir un critère (une condition nécessaire et suffisante) purement algébrique permettant de constater que toutes les équations et inéquations (considérées) possèdent au moins une solution réelle commune (racine). » Voir également le commentaire du lemme 2.5, p. 225-226.

<sup>(63)</sup> *Ibid.*, 2.1, p. 216.

<sup>(64)</sup> Tarski connaît d'abord le théorème de Sturm par les exposés de H. Weber, de C. Runge et de B. L. Van der Waerden. Ensuite, la note 12 de *A Decision Method...* permet de le penser, il a lu le mémoire original de l'auteur.

<sup>(65)</sup> Il est intéressant de remarquer que Tarski et Sylvester, parlant tous deux de la nécessité d'un critère algébrique de la réalité des racines d'un polynôme, aboutissent, *via* des intérêts fort différents, à des solutions relevant de points de vue totalement distincts.

trop loin, disons que, dans le système de Tarski, un énoncé sans quantificateur est une combinaison (booléenne) de formules atomiques du type :

$$\alpha = \beta \quad \text{ou} \quad \alpha > \beta,$$

où  $\alpha$  et  $\beta$  sont des « termes » algébriques, c'est-à-dire des expressions logiquement bien formées à partir des éléments que sont les symboles de constantes ou de variables parcourant l'ensemble des nombres réels (0, 1,  $x$ ,  $y$ ,  $z$ , ...), et les symboles des opérations algébriques élémentaires, ce mot étant pris au sens usuel : (+, —,  $\times$ , :) <sup>(66)</sup>. Un énoncé sans quantificateur du système logique élémentaire (au sens de Tarski) pour le corps commutatif, ordonné, des nombres réels, c'est donc précisément une combinaison, ou, comme on dit habituellement, un système d'équations et d'inéquations algébriques. On voit le rapport, essentiel, entre algébrique tel que l'entend spontanément un mathématicien, et tel que l'entend aussi Tarski parlant de « critère purement algébrique », et « énoncé sans quantificateur », notion logique forgée dans le contexte des solutions partielles au problème de la décidabilité <sup>(67)</sup>. Généralisant le théorème de Sturm et rendant visible la procédure algébrique et la procédure logique impliquées dans cette généralisation, Tarski a montré que la théorie élémentaire du corps ordonné des nombres réels admet l'élimination des quantificateurs, c'est-à-dire équivaut à une théorie algébrique en un sens qui rejoint l'acception mathématique traditionnelle.

Celle-ci exige pour le moins, en effet, l'exclusion, dans les raisonnements, des notions de continuité ou d'infiniment petit, et recoupe ainsi l'opposition banale Algèbre/Analyse ou théorie

<sup>(66)</sup> Tarski ne se sert en fait que des symboles des trois opérations d'addition, de soustraction et de multiplication. Cela implique une généralisation immédiate de ses résultats à la structure d'anneau unitaire ordonné. Insistons ici sur le fait que des symboles comme celui de l'exponentielle, du logarithme ou de l'intégrale sont exclus de ce langage.

<sup>(67)</sup> Voir la liste de ces résultats jusqu'à 1951 dans l'introduction à *A Decision Method for Elementary Algebra and Geometry*, p. 12. Y ajouter celui de J. Herbrand, *Recherches sur la théorie de la démonstration*, éd. Jean Van Heijenoort (Paris : PUF, 1968), chap. IV. Une théorie T, formulée dans un langage L, est décidable s'il existe une procédure, dite de décision, permettant de décider dans chaque cas particulier si une proposition, écrite dans L, peut être démontrée — se souvenir qu'une démonstration est une suite finie de formules ! — par des moyens internes à T (voir par exemple Tarski, *A general method in proofs of undecidability*, dans *Undecidable Theories*, en collab. avec Mostowski et R. M. Robinson (Amsterdam : North-Holland, 1953), p. 3).

des équations/théorie des fonctions. Donner une théorie de  $\mathbf{R}$ , qui fût sinon algébrique du moins non analytique (ou non géométrique), a donc signifié à la fin du XIX<sup>e</sup> siècle : trouver un ensemble d'axiomes, ne comportant parmi eux aucun axiome de continuité, et caractérisant néanmoins  $\mathbf{R}$ , en ce sens que tout théorème vrai dans  $\mathbf{R}$  soit déductible de cet ensemble. C'est ainsi que Dedekind a réussi à construire l'ensemble ordonné des nombres réels grâce à la méthode des coupures <sup>(68)</sup> ; que Hilbert, qui voulait échapper à une expression directe de la continuité, a dû cependant adjoindre aux axiomes de *corps* ordonné l'axiome d'Archimède et un « axiome d'inextensibilité » <sup>(69)</sup> et que les mathématiciens de l'école algébrique américaine ont cherché quantité de variantes au système de Hilbert <sup>(70)</sup>. Tarski connaissait évidemment tous ces travaux <sup>(71)</sup>. Lui-même choisit tantôt l'axiome de Dedekind, comme dans son *Introduction à la logique* <sup>(72)</sup>, tantôt le théorème de la borne inférieure : axiome XVII du « système élémentaire d'algèbre des réels » construit dans « La complétude de l'algèbre et de la géométrie élémentaires », tantôt et de préférence, une version algébrique du théorème des valeurs intermédiaires : axiome XVII' du même système <sup>(73)</sup>. Cette préférence est par elle-même significative : il était bien connu que le contenu de l'axiome XVII' était au fondement d'une théorie des équations construite sans emprunt extérieur de concept ou de méthode.

Il est donc bien évident que Tarski veut construire un sys-

<sup>(68)</sup> Dedekind, *Stetigkeit und irrationale Zahlen*, 1<sup>re</sup> éd. (Braunschweig, 1872) ; trad. franç. : *La bibliothèque d'Ornicar ?* (Diffusion Paris : Le Seuil, 1978).

<sup>(69)</sup> Hilbert, *Über den Zahlbegriff*, *Jahresber. der Deutsch. Math.-Vereinigung*, 8 (1900), 180-184.

<sup>(70)</sup> Ces travaux, dus notamment à E. H. Moore, E. V. Huntington et O. Veblen, se trouvent dans les premiers volumes (en particulier le VI et le VII) des *Transactions of the American Mathematical Society*.

<sup>(71)</sup> Le chapitre X de son *Introduction à la logique* (1<sup>re</sup> éd. en polonais, 1936 ; trad. franç., Paris : Gauthier-Villars, 1960) intitulé : « Fondements de l'arithmétique des nombres réels », en témoigne. Voir en particulier la n. 1 de la p. 191 où Tarski mentionne, à côté du système de Hilbert, celui, ou plutôt ceux de Huntington (dont il recommande par ailleurs la monographie : *The Fundamental Propositions of Algebra*, éd. by J. W. A. Young (New York, 1911)). Dans « La complétude de l'algèbre... » la note de la p. 232 renvoie au système de Huntington pour les complexes. Dans *Les ensembles définissables de nombres réels* (1931), traduit dans *Logique, sémantique, métamathématique*, t. I, il renvoie à un système de O. Veblen.

<sup>(72)</sup> Chap. X, trad. franç., p. 191-198.

<sup>(73)</sup> Cet axiome est celui conservé par la version de 1948 : *A Decision Method...*, n. 9, p. 48-49.

tème algébrique des nombres réels <sup>(74)</sup>. Or il apporte trois résultats nouveaux qui sont, au fond, des déterminations concrètes du projet algébrique d'ensemble :

1 / S'apercevant que tous les axiomes de continuité sont logiquement complexes puisqu'ils exigent l'usage de variables désignant des ensembles de nombres réels, il montre néanmoins qu'on peut caractériser dans le langage du calcul des prédicats du premier ordre la théorie algébrique des nombres réels. D'où le qualificatif « élémentaire » pour son système (infini puisque XVII' est un schéma d'axiome).

2 / En explicitant au préalable la collection des symboles primitifs du langage utilisé, il met en évidence le matériau nécessaire : deux symboles pour les constantes d'individus 0 et 1 ; des symboles de variables d'individus <sup>(75)</sup> ; trois symboles pour les opérations d'addition, de soustraction et de multiplication ; enfin deux symboles pour les relations d'égalité et d'inégalité. Rien que de très élémentaire en somme (au sens banal du terme), mais qui suffit cependant à construire l'objet « polynôme à coefficients réels », et à ramener la question métamathématique de la décidabilité de la théorie élémentaire de  $\mathbf{R}$  à celle de la résolubilité dans  $\mathbf{R}$  d'un système d'équations et d'inéquations.

3 / La « méthode de Sturm généralisée » permet de montrer que la théorie algébrique élémentaire <sup>(76)</sup> de  $\mathbf{R}$  admet l'élimination

<sup>(74)</sup> Nous laissons ici de côté la construction algébrique des corps réels d'Artin-Schreier, publiée en 1927, mais dont Tarski ne tient explicitement compte qu'en 1948 : *A Decision Method...*, n. 9 et 15.

L'axiomatique d'Artin-Schreier est celle qui est en usage aujourd'hui. Nous en avons analysé ailleurs la signification et la portée (Communication au Colloque d'Histoire des Mathématiques, CRM, Marseille-Luminy, juin 1983) ; voir aussi l'ouvrage (à paraître) : *Corps et modèles. Aspects de la construction de l'algèbre réelle*, II<sup>e</sup> partie.

<sup>(75)</sup> Ces variables prennent leurs valeurs dans le corps des nombres réels ou, comme le remarque ensuite Tarski, dans le corps des nombres réels algébriques. C'est pourquoi il s'agit bien d'une théorie *algébrique* de  $\mathbf{R}$ .

<sup>(76)</sup> C'est à dessein que nous parlons de théorie algébrique et élémentaire, de même que Tarski écrit : « système élémentaire d'algèbre des réels ». Aujourd'hui, compte tenu de l'observation que les structures de l'algèbre abstraite : groupe, anneau, corps, module, etc., et leurs différentes spécifications comme, par exemple, groupe abélien divisible, corps algébriquement clos, anneau ordonné, etc., peuvent être axiomatisées (de façon finie ou dénombrable) dans le langage logique du calcul des prédicats du premier ordre, certains logiciens tendent à considérer qu'une théorie est caractérisable algébriquement si elle est caractérisable dans le premier ordre. C'est par exemple, la convention adoptée par Kreisel-Krivine, *Eléments de logique mathématique*, Mono-

des quantificateurs et qu'elle est, par conséquent, décidable. Compte tenu de 1 / et 2 /, on comprend l'importance qu'accordait Tarski à ce que le théorème de Sturm permit d'avoir un « critère purement algébrique » de la réalité des racines d'un polynôme <sup>(77)</sup>.

Nous venons de montrer que la question de la nature, algébrique ou non, du théorème de Sturm se pose surtout depuis l'existence de la méthode d'élimination des quantificateurs de Tarski (qui est la « méthode de Sturm généralisée »). Indice d'une discussion encore actuelle sur l'utilisation de méthodes de démonstration non algébriques pour des propositions exprimables algébriquement, elle correspond aussi à un intérêt métamathématique aigu pour la signification du terme « algébrique ». Cet intérêt est thématiqué notamment dans certains travaux de Tarski <sup>(78)</sup> et d'A. Robinson <sup>(79)</sup> que nous ne pouvons examiner ici.

## 2 / Le premier exemple d'élimination des quantificateurs pour le cas réel

Ce qu'a fait Tarski du théorème de Sturm justifie largement la « sensation » produite à sa naissance. Tous ces caractères que lui reconnaissaient, par une vue plus ou moins claire, les mathématiciens, le logicien les a confirmés : simplicité, effectivité algorithmique (de principe), importance fondamentale pour les mathématiques.

graphie de la SMF (Paris : Dunod, 1967), chap. 3, résumé, p. 33, et appendice I, p. 150. Mais cette convention est loin d'être acceptée d'une façon générale.

Signalons qu'en tout état de cause une théorie algébrique n'admet pas forcément l'élimination des quantificateurs. On a ainsi montré récemment qu'un anneau ordonné n'a cette propriété que s'il vérifie les axiomes de corps ordonné et réel-clos.

<sup>(77)</sup> Dans la deuxième version de son travail, Tarski insistera plutôt sur le fait que le théorème de Sturm fournit une condition élémentaire (du premier ordre) de l'existence d'une racine réelle pour un polynôme donné. Voir *A Decision Method...*, n. 12, p. 51, 2<sup>e</sup> al.

<sup>(78)</sup> Citons comme un bon représentant de ce vecteur dans les recherches de Tarski : *Equational Logic and equational theories of algebra*, dans *Coll. Log. Found. Math.* Hannover, 1966, publié en 1968 (Amsterdam : North-Holland), 275-288, qui reprend certaines idées exprimées par J. C. C. McKinsey dans *The decision problem for some classes of sentences without quantifiers*, *Journal of Symbolic Logic*, 8 (1943).

<sup>(79)</sup> Pour A. Robinson, une « algèbre d'axiomes » est un ensemble d'axiomes, qui comporte un symbole de relation pour l'égalité. Une structure est algébrique si elle est axiomatisée par une algèbre d'axiomes (voir par exemple *On the Metamathematics of Algebra* (Amsterdam : North-Holland, 1951), chap. VI). Robinson définit par ailleurs la notion métamathématique de prédicat (formule) algébrique par analogie avec celle d'élément algébrique sur un corps (*ibid.*, chap. VIII).

Que d'auteurs ont été éblouis par la simplicité de notre théorème <sup>(80)</sup> ! Tarski a montré d'une façon précise qu'il se laisse en fait algébriser sans problème et sans secours extérieur <sup>(81)</sup> ; et qu'il s'écrit dans un langage logique du premier ordre. Bref, tout comme le système construit par Tarski pour les nombres réels, et dans lequel on peut l'exprimer, il est algébrique et élémentaire.

Tarski a encore confirmé l'intention, diversement mais très fréquemment formulée dès 1829, qu'il s'agissait d'un théorème capital. Il a, en fait, déterminé la raison de cette importance : le théorème de Sturm permet de faire, pour le cas du corps ordonné  $\mathbf{R}$  (muni, outre de la relation  $=$ , de la relation  $>$ ), une théorie parallèle à celle de l'élimination algébrique classique. — On appelle ainsi l'ensemble des méthodes de résolution des systèmes d'équations algébriques à plusieurs inconnues (systèmes qui, du point de vue logique, sont écrits avec la seule relation  $=$ ) par élimination successive des inconnues. — Il apparaîtra, ensuite, plus clairement que cette théorie peut être, dans certains cas, avantageusement substituée à celle de l'élimination algébrique ou à d'autres méthodes algébriques particulières, sa généralité faisant d'elle une méthode simple et d'une grande clarté conceptuelle. A. Robinson en a ainsi déduit une preuve triviale du dix-septième problème de Hilbert <sup>(82)</sup>.

Tarski, enfin, a tiré tout le parti théorique du caractère *effectif* du théorème. La tradition logique, représentée par G. Boole et par E. Schröder, d'une algèbre de la logique, c'est-à-dire d'une théorie logique à l'*image* de la théorie des équations, fut à l'origine de l'idée que l'élimination des quantificateurs d'une formule logique rendait soluble pour cette formule le problème de la décision. C'est ainsi qu'on trouve dans un mémoire de L. Löwenheim (1915), grandement influencé par les travaux et le langage de Schröder, la première esquisse de cette méthode de décision appli-

<sup>(80)</sup> Nous avons cité plus haut Terquem (référence dans la n. 3). Mais nous pourrions évoquer bien d'autres éloges, de Liouville, de Borchardt, de Serret, de Hermite, etc. Retenons peut-être que Hermite avait salué dans ce théorème « un exemple rare de simplicité et d'élégance », qui « ouvre l'ère nouvelle de l'Algèbre moderne » (lettre à Borchardt du 28 janvier 1854, dans *Oeuvres*, t. I).

<sup>(81)</sup> Tout en étant moins général que « l'Algèbre réelle » d'Artin-Schreier, le système élémentaire d'algèbre des réels de Tarski appréhende cependant le théorème de Sturm d'une façon analogue (et fort éloignée de l'approche également algébrique de Sylvester ou d'Hermite).

<sup>(82)</sup> On ordered fields and definite functions, *Mathematische Annalen*, 130 (1955).



quée au calcul monadique <sup>(83)</sup> des prédicats du premier ordre. D'autre part, lorsque Hilbert et Ackermann formulent en 1928 <sup>(84)</sup> le problème général de la décision, le paradigme de la théorie des équations est bien entendu toujours présent, mais un accent particulier est mis sur l'aspect de calculabilité effective. Pour eux, en effet, l'usage du symbolisme logique pour l'exposé formel des théories (axiomatiques en particulier) laisse espérer la mise au point d'un traitement permettant de décider, systématiquement (c'est-à-dire pour une formule quelconque) et en un nombre fini d'étapes, si telle formule donnée est valide dans telle théorie. En termes plus modernes, on dirait que c'était définir par là l'idée d'un algorithme de décision <sup>(85)</sup>.

Or, Tarski remarque (en 1948) que les mathématiques traditionnelles offrent au moins deux exemples de tels algorithmes. Celui d'Euclide fournit, en effet, une procédure de décision pour la classe des propositions vraies du type «  $p$  et  $q$  sont premiers entre eux », où  $p$  et  $q$  sont des nombres entiers ou des polynômes à coefficients constants. Celui de Sturm permet de décider la vérité de propositions de la forme « le polynôme  $p$  a exactement  $n$  racines distinctes dans l'intervalle  $(A, B)$  ». Par construction, mais aussi par nature, le second est plus complexe que le premier, car les propositions décidées par lui nécessitent normalement pour leur formulation logique l'usage de quantificateurs. Si l'on explicite, on trouve une formule du genre :

$$\begin{aligned} \exists x_1 \exists x_2 \dots \exists x_n \{ & [(x_1 < x_2) \wedge (x_2 < x_3) \wedge \dots \wedge (x_{n-1} < x_n) \\ & \wedge (A \leq x_1) \wedge (x_n \leq B) \\ & \wedge (p(x_1) = 0) \wedge (p(x_2) = 0) \wedge \dots \wedge (p(x_n) = 0)] \\ & \Rightarrow \forall y [(A \leq y \leq B) \wedge (p(y) = 0)] \\ & \Rightarrow ((y = x_1) \vee (y = x_2) \vee \dots \vee (y = x_n)) \} \end{aligned} \quad (86).$$

<sup>(83)</sup> Monadique signifie qu'on se restreint aux prédicats à une place, ce qui implique quantification non multiple.

<sup>(84)</sup> *Grundzüge der theoretischen Logik* Berlin : Springer-Verlag, 1938), § 11-12, p. 72-81.

<sup>(85)</sup> Voir par exemple la définition des concepts de décidabilité et de calculabilité dans H. Hermes, *Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit* (Berlin : Springer-Verlag, 1961), chap. 1.

Un ensemble d'éléments  $E$  est décidable si et seulement si sa fonction caractéristique  $\chi_E$  est calculable. Une fonction est calculable s'il existe un algorithme, c'est-à-dire une procédure de calcul effectif, pour calculer la valeur de la fonction pour n'importe quelle valeur de son argument.

<sup>(86)</sup> «  $A \leq x_1$  » est une abréviation pour : «  $(A < x_1) \vee A = x_1$  ».

Cependant, Tarski montre que, grâce au théorème de Sturm, cette proposition est équivalente à une proposition *sans* quantificateurs. En effet, on a vu que  $n = N_A - N_B$  <sup>(87)</sup> ; or cette égalité peut être exprimée par une combinaison booléenne d'égalités et d'inégalités polynomiales dont les termes sont obtenus rationnellement à partir des coefficients de  $p$  et des bornes de l'intervalle. Par exemple, dans le cas d'un polynôme du second degré  $x^2 + bx + c$  la suite de Sturm est  $\left(x^2 + bx + c, 2x + b, \frac{b^2}{4} - c\right)$  ;  $N_A - N_B = 2$  (en supposant qu'on est dans ce cas) s'exprime par la conjonction <sup>(88)</sup> :

$$(b^2 - 4c > 0) \wedge (A^2 + Ab + c > 0) \wedge (2A + b > 0) \\ \wedge (B^2 + bB + c > 0) \wedge (2B + b > 0).$$

Dans le cas où l'intervalle n'est pas spécifié, on retrouve la condition bien connue :  $b^2 - 4c > 0$ .

Ainsi, le théorème de Sturm est un magnifique exemple d'élimination (implicite) des quantificateurs. C'est même le premier exemple non trivial : on ne sait pas *a priori* si un polynôme a une racine *réelle* dans un intervalle donné, alors que si on se place, par exemple, dans le corps des complexes on sera sûr de toute façon de l'existence de racines en nombre égal au degré du polynôme.

Sturm aurait sans doute été bien surpris de la métamorphose de son théorème. L'historien l'est, à vrai dire, à peine moins. Ne voilà-t-il pas, en effet, un résultat de l'Analyse, dont la rigueur et le progrès furent étroitement liés, à la fin du siècle dernier, à la théorie de la quantification multiple (G. Frege), qui se voit ensuite transformé en prototype de méthode d'*élimination* des quantificateurs ? On dira pour le moins que les voies de l'histoire sont parfois bien indirectes !...

CNRS, Paris.

Hourya BENIS-SINACEUR.

<sup>(87)</sup> *Supra*, 1<sup>er</sup> paragraphe, p. 102.

<sup>(88)</sup> Ce cas volontairement choisi pour sa simplicité ne donne cependant pas une dée exacte de la longueur de l'algorithme pour un polynôme de degré plus élevé.